DOCUMENT RESUME

ED 441 763		SP 039 219

AUTHOR	Craven, Rhonda G.; Marsh, Herbert W.; Debus, Raymond L.
TITLE	Teacher-Administered Self-Concept Interventions: Do Diffusion Effects Exist?
PUB DATE	2000-04-24
NOTE	19p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Control Groups; *Diffusion (Communication); Elementary Education; Elementary School Students; Experimental Groups; Foreign Countries; *Research Methodology; *Self Concept Measures; Student Attitudes
IDENTIFIERS	Australia (Sydney)

ABSTRACT

Within-class experimental designs (with experimental and control groups in the same classroom) are subject to diffusion effects whereby both experimental and control students benefit from the intervention, thereby contaminating the control group and biasing evaluations of intervention effects. In support of diffusion effects, this study shows that a classroom intervention resulted in systematically higher academic self-concepts for internal (within class) controls compared to external (between class) control groups. The construct validity of the interpretation of this difference as a diffusion effect was supported by observer and teacher comments and ratings of teacher success in focusing the intervention on experimental students and different patterns of results for teachers who were more or less successful in maintaining this focus. Potential dangers in sole reliance on internal within-class control groups may outweigh advantages of this expedient experimental design. (Contains 19 references.) (Author/SM)

Teacher-Administered Self-Concept Interventions: Do Diffusion Effects Exist?

Rhonda G. Craven, Herbert W. Marsh
University of Western Sydney, and
Raymond L. Debus, University of Sydney

This research was based on a PhD thesis by the first author and funded, in part, by a grant from the Australian Research Council. We thank Lawrence Roche for his assistance in conducting the research and helpful comments on earlier versions of this paper. We would also like to thank the principals, teachers, and students who made the research possible.
Requests for further information about this investigation should be sent to Associate-Professor Rhonda Craven, Deputy Director, SELF Research Centre, University of Western Sydney, Macarthur, PO Box 555, Campbelltown, New South Wales, Australia, 2560. Tel: (02) 9772-6657 FAX (02) 9772-1565 E-mail: r.craven@uws.edu.au.

Abstract

Within-class experimental designs (with experimental and control groups in the same classroom) are subject to diffusion effects whereby both experimental and control students benefit from the intervention, thereby contaminating the control group and biasing evaluations of intervention effects. In support of diffusion effects, we show that a classroom intervention resulted in systematically higher academic self-concepts for internal (within-class) controls compared to external (between class) control groups. The construct validity of the interpretation of this difference as a diffusion effect was supported by observer and teacher comments and ratings of teacher success in focusing the intervention on experimental students, and different patterns of results for teachers who were more or less successful in maintaining this focus. Potential dangers in sole reliance on internal within-class control groups may outweigh advantages of this expedient experimental design.

In completely within-classroom experimental designs, the entire experimental design (i.e., all experimental and control groups) is replicated within each classroom. Typically, these designs involve matching students in each class and then randomly assigning each matched student either to an experimental group that receives a teacher-administered intervention or to an internal (within-class) control group. Such designs are very efficient in that they can be implemented with a small number of classes or even a single class and provide more precise estimates of the intervention effects. Within-class designs differ from between-classroom designs in which all students within a given classroom are in the same condition such that no experimental students are in the same classroom as the external (between-class) control students. Although there are many variations of these basic experimental designs, the distinguishing feature is that random assignment is conducted at the level of the individual student for within-classroom designs but at the level of the classroom for between-classroom designs. This distinguishing feature has important implications for the design, analysis, and interpretation of classroom research. For purposes of the present investigation we have selected a component of a larger, previous study (Craven, 1996) in which to demonstrate potential biases produced by within-class designs. Because the purpose of this study is methodological, our focus is on issues of design, analysis, and interpretation of results based on within-class designs.

Potential Threats to Internal Validity Produced by Within-classroom Designs

Diffusion Effects

In their classic discussion of threats to internal validity, Cook and Campbell (1979) discuss a number of ways in which direct or indirect interaction between experimental and control groups can invalidate comparisons between these groups. They caution that diffusion or imitation of treatments can occur "when treatments involve informational programs and when the various experimental (and control) groups can communicate with each other, respondents in one treatment group may learn the information intended for others" (p. 54). They also suggested that this problem is particularly acute in quasi-experimental designs that attempt to ensure that control and experimental groups are similar, and include a physical closeness of such groups so that they can communicate. Good and Brophy (1977) have described this phenomenon as a treatment that radiates to nontarget participants. More recently researchers have described this issue as "leakage" (Plewis and Hurry, 1998) and Craven (1996) specifically referred to this phenomenon as what she termed a "diffusion effect". Given the latter term is consistent with previous and current researchers' descriptions of the issue, throughout this paper we will refer to this threat to internal validity as a diffusion effect.

Although listed as threats to internal validity that are distinct from diffusion effects, Cook and Campbell (1979) listed other potential threats that may be relevant to evaluating results for internal (within-class) comparison groups: **compensatory equalization** (providing additional benefits to control group participant to compensate for benefits lost by not being in the experimental group); **compensatory rivalry** (control group participants trying harder to compensate for the expected difference in favor of the experimental group), and **resentful demoralization** (control group participants giving up or not trying as hard because they are demoralized about not receiving the benefits of the intervention). Whereas diffusion effects are typically assumed to reduce the size of the intended effects of an intervention compared to a design in which the effects were not contaminated by diffusion effects, other threats to internal validity such as resentful demoralization could actually increase the size of the effects. Furthermore, these various threats are not mutually exclusive so that it is difficult to anticipate their net effect. For example, some control-group participants might try harder (compensatory rivalry) whereas others might not try as hard (resentful demoralization).

A diffusion effect may be present in within-classroom designs when teachers are asked to deliver the intervention to target students in the experimental groups and not to deliver the intervention to nontarget students in the control groups (internal, within-class control groups). Hence, the internal validity of within-class designs may be weaker than between-class (or school) designs in which the external control groups have no interaction or awareness of the experimental group. A critical issue is whether teachers are able and willing to deny control students the potential benefits of the intervention. Even with careful training and monitoring, teachers are likely to differ in the fidelity with which they implement the intervention. Hence, within-class designs are vulnerable to diffusion effects whereby the teacher-mediated intervention diffuses to control group participants. For example, teachers might incorporate changes associated with the intervention into their natural teaching repertoires if they deem the new strategy as potentially successful in inducing positive changes in student behavior. If this did occur, then this change in teacher behavior may diffuse to nontarget control students and result in corruption of the within-class control group. Even if teachers do not apply experimental procedures to nontarget students, it is also possible for students to experience a teacher-mediated intervention vicariously in that they may hear target students receiving feedback and use this feedback as a basis of altering their future behavior (Bandura, 1986).

Diffusion effects are problematic in that the contamination of the internal (within-class) control group may result in biased estimates of the intervention effect. Hence, if diffusion effects are present, results are difficult to interpret in that internal validity has been compromised and the control group may have been influenced directly or indirectly from aspects of the intervention. Despite the potential bias of such effects for teacher-mediated interventions, researchers generally overlook the possibility that a treatment has inadvertently affected control students. Yet "if the classroom ecology is to be disturbed, it is important to assess how changes in teacher behavior affect <u>all</u> students" (Good and Brophy, 1974, p. 391). Therefore the possibility that teacher-mediated treatments could diffuse to the control group needs to be explicitly examined in teacher-mediated intervention studies based on a within-class experimental design.

<u>Research Evidence For Diffusion Effects.</u>

Several studies have indicated that diffusion effects may be present. Withall (1956), in an early study designed to examine teacher's classroom interactions, advised a teacher that 8 specific students could benefit from more teacher interaction. Based on classroom observations, Withall found that the teacher increased his interaction with target students but teachers' interactions with nontarget students also rose significantly. The results of the Withall study suggest that a diffusion effect was present in that the teacher changed his behavior towards all students not just solely target students.

Good and Brophy (1974) explored whether feedback given to teachers could change teacher behavior towards target students and observed the effects of changes in teacher behavior for both target and nontarget students. The training was an interview with individual teachers to make them aware of negative interactions with target students in comparison to positive interactions with nontarget groups. Seven of the eight participating teachers showed large changes in the pattern of their interactions with target students. Whereas 4 of these 7 restricted the intervention to target students, three teachers also changed the pattern of their interaction with nontarget students. Good and Brophy (1974, p. 404) concluded, "when teachers did change their behavior toward target children, they also tended to change their behavior (in the same direction) toward nontarget children". This diffusion to nontarget students benefited the nontarget students, but contaminated comparisons between control and intervention students and negatively biased estimates of the intervention effect based on such comparisons. Clarke and Cornish (1972) reported a similar effect in a criminological study. They found that staff of a penal institution for teenage boys implemented a "therapeutic community" intervention to both experimental and control groups rather than solely utilizing existing orthodox methods with the control group. Thereby staff in this study changed their behavior toward nontarget children so that the intervention for the control group became more like the intervention group over time.

Diffusion effects can result in the increase of desirable teacher behaviors for control as well as experimental students. It is, however, also likely that a diffusion effect could result in a decrease of the frequency of undesirable teacher behaviors to students in both experimental and control groups. For example, Cooper (1977) asked teachers to refrain from criticizing experimental participants after the student initiated an interaction. Observation four weeks after the intervention revealed that teachers had stopped criticizing all students, not just students assigned to the experimental group.

Meta-analyses of intervention studies have also identified the presence of positive changes to control groups although these apparently have not been attributed to diffusion effects. For example, Hattie's (1992, p. 227) meta-analysis of self-concept enhancement studies found that there was an effect for positive change in control groups with an average effect size of .12 based on 51 effect-sizes, suggesting that this could be explained by Hawthorne, maturation, experimenter, or "copy cat" effects. Although she did not refer specifically to diffusion effects nor categorize these effects according to specific types of research design, these results may be suggestive that such effects may be present in some of the studies.

Other Potential Biases Associated With Within-class Designs.

Diffusion effects imply a bias such that control group students benefit from an intervention that is supposed to benefit only experimental group students. As noted earlier, one possible mechanism whereby this might take place is through vicarious reinforcement in which teacher praise not only has the desired effect on target students who receive the reinforcement, but also has similar effects on nontarget students who merely observe this process (Bandura, 1986; Sharpley, 1985). Thus, for example, the nontarget students may assume that they will be praised by the teacher in the future for such behavior or even reinforce themselves when they perform the desired behavior in the future. If, however, two students are concurrently performing the same behaviors and the teacher explicitly reinforces only a target student, then the predicted effects on the nontarget student are more complicated (Sharpley, 1985). The effect on nontarget students may be positive due to a vicarious reinforcement effect or the consequences of the behavior other than teacher reinforcement.

Conversely the nonreward of control participants, may also extinguish the desired behavior for nontarget students and thus have the opposite effect. Bandura (1986, p.286), for example, suggests that "those whose efforts go unrecognized are more likely to be disheartened than inspired by seeing others receiving recognition to which they also feel entitled" (also see Cook and Campbell, 1979, for discussion of 'resentful demoralization'). In a classic illustration of this negative implicit reward effect, Sechrest (1963) studied pairs of students who concurrently completed two different puzzles. One student in each pair -- the target student -- was praised or criticized whereas the nontarget (internal control) student in each pair received no reinforcement. An additional external control group completed the puzzles alone and received neither praise nor criticism. Praise led to better subsequent performance for target students who received the praise, but poorer performance for the nontarget students who merely observed the praise of target students. Conversely, criticism led to poorer performance for target students, but better performance by the nontarget pair who merely observed target students being criticized. In a review of implicit rewards in classroom settings, Sharpley (1985) emphasized that the use of implicit rewards can lead to poorer performances when these students have previously been rewarded for the same behaviors.

A multilevel perspective.

Selection of the appropriate unit of analysis – the individual student or the classroom – is an important issue in classroom research that is particularly relevant to the discussion of diffusion effects. Because students are clustered within classrooms, there are multiple levels of analysis. The lack of independence among students within the same class – a clustering effect – is likely to invalidate the assumptions underlying traditional statistical analyses based on response by individual students. Particularly for within-class designs in which the same teacher administers the intervention to experimental students and does not administer the intervention to control students, the extent of diffusion effects is likely to vary from teacher to teacher. When there are multiple classrooms, researchers can use multilevel modeling (e.g., Bryk & Raudenbush, 1992; Goldstein, 1995), to appropriately incorporate both the individual student and the classroom into the same analysis.

Plewis and Hurry (1998) provided a technical discussion of statistical models and the conduct of multilevel analysis, indicating why a multilevel perspective is important to consider in the design and analysis of classroom intervention studies. In their demonstration, they pursued further analyses of selected components of a reading intervention. In this research, children with reading difficulties received individual tutorial sessions from a trained reading recovery teacher. In each of the 22 schools implementing the program, the six poorest readers were identified, 3 or 4 of these students were allocated to the intervention in which they were withdrawn from class to receive individual instruction, and the remaining students were allocated to a within-school control group. The comparison of results for these two groups suggested a positive effect of the intervention using an appropriate standard error based on differences between classrooms. They emphasized that, as is typically the case, the standard error based on analyses of individual students (that ignores the classroom level) was substantially smaller than the more appropriate standard error based on multilevel modeling, providing a positively biased test of the statistical significance of the intervention effect. More generally, the size of the standard error in multilevel modeling can vary in size from the typically smaller standard error based on analyses of large numbers of individual students (i.e., each student is considered to be a separate case) to the typically larger standard error based on analyses class mean (i.e., each class is considered to be a separate case). Where it falls along this continuum depends on the size of the clustering effect (the extent to which students within each class are more similar to each other than to students in different classes).

In discussing potential limitations of this particular reading intervention study, Plewis and Hurry emphasized the possibility of diffusion effects (which they refer to as "leakage"). They suggested that internal control groups in their study could be affected by diffusion effects in that all teachers in experimental schools -- not just teachers implementing the intervention -- were trained in reading recovery and that the trained reading recovery teachers may also have taught control students utilizing the intervention methods when in the role of regular classroom teacher. This study provided a potential test for diffusion effects in that a further 41 control schools were selected by local education authorities based on similar student intake to experimental schools (also see discussion of this approach by Craven, 1996). Although these external control schools were not randomly assigned to conditions, the six poorest readers were assigned to an external control group. Plewis and Hurry conducted separate multilevel analyses based on the internal (within school) and external (between school) control groups. Although both analyses showed statistically significant intervention effects, effects based on the external control group were slightly larger (an effect size of .79 vs. .62). Plewis and Hurry interpreted this difference to be "consistent with some leakage" (p. 22). They did not, however, provide a strategy for testing the statistical significance of this apparent difference that may be modest in relation to probable sampling error. Furthermore, because the external control groups were not based on random assignment, they cautioned that even the small observed differences "may be confounded with other design differences, particularly in allocation" (p. 22). They concluded, however, that "because of possible leakage, in this example the classic comparison between intervention and control children in the same school is a demanding one as far as demonstrating an intervention effect is concerned" (p. 22). Because the focus of their study was on the application of multilevel modeling rather than diffusion effects per se, and the results were statistically significant for both analyses, the authors did not pursue more formal comparisons of the internal and external control groups. Rather their main message was that "Whatever design is adopted to the effectiveness of an educational intervention, this paper has shown that the analysis of the data so generated needs to be located within a multilevel framework" (p. 24).

The Present Investigation

The research reviewed here implies that diffusion effects may confound the interpretation of intervention studies. However, whilst researchers may be aware of and imply that this issue is important (often in passing), few researchers have conducted rigorous tests of diffusion effects using appropriate research designs and statistical analyses. Researchers tend to focus on intervention effects for target participants rather rigorously analyzing and interpreting the impact of interventions on intended control participants. As noted previously, in educational environments "it is important to assess *(emphasis added)* how changes in teacher behavior affect all students" (Good and Brophy, 1974, p. 391), and "predicting and controlling for such effects should be of special concern to those who propose to change teacher behavior" (Good and Brophy, 1974, p. 405, *emphasis added*). Unfortunately, most research demonstrations of diffusion effects are anecdotal in nature, lacking rigorous research designs and appropriate statistical analyses to better understand these effects that would serve to strengthen intervention design, implementation and evaluation procedures. Hence, given the importance of teacher-mediated interventions for educational research, and the problematic nature of potential diffusion effects for data analysis and interpretation, it is important to better document the occurrence, causes, and research implications of such effects.

The purpose of the present investigation was to provide a basis for better informing research pedagogy in relation to these issues. In so doing, we demonstrate useful methodological approaches for investigating diffusion effects and fully exploring the implications of the procedures and findings of this study for future practice. Specifically, the key purposes of the investigation were to: a) Highlight and emphasize the importance of diffusion effects as a substantive methodological issue; b) Provide an overview of the methodology employed for a model study to identify useful methodological approaches; and c) explore the implications of the findings to strengthen future research.

The model study was based on a large-scale within-class self-concept intervention (Craven, 1996) delivered by teachers. Important methodological features of the study include: a) predicting and controlling for possible diffusion effects in the research design as suggested by Good and Brophy (1974); b) focusing on assessing how changes in teacher behavior affect nontarget students; c) conducting appropriate tests of statistical significance to demonstrate the presence of diffusion effects; d) utilizing a strong research design incorporating a randomly assigned within-class control group and randomly assigned within-school external control group based on matching procedures to test for diffusion effects; and e) utilizing a synergetic blend of quantitative and qualitative research methods to assist in illuminating the presence and operation of diffusion effects.

Method

Participants

Participants for the self-concept intervention study were 1557 students aged from 8 to 10, from 50 classes in 8 schools in metropolitan Western Sydney, Australia. Pretest total academic self-concept scores measured by the Self-Description Questionnaire I (SDQ-I; Marsh, 1990) were used as the criterion for selecting students to participate in the study and matching students who were then randomly allocated within each class to experimental, and internal-control (within-class) groups. One class in each school was randomly assigned to the external-control group after randomly selecting the year group to be allocated to the external control group in each school. From each of the 50 classes participating in the study, 18 students with the lowest total academic self-concept scores were selected to participate from an average of 30 children per class.

The 18 students from each of 42 experimental classes receiving the intervention were matched in triplicates by sex, age and level of academic self-concept. The matched participants were then randomly assigned to control or experimental interventions. This resulted in one participant being assigned to the internal control group ($\underline{N}$ = 252 across all 42 classes) and the remaining two participants being assigned to the experimental interventions that focused on enhancing math and verbal self-concepts through interventions delivered either by the teacher or by research assistants. One additional class from each of the eight participating schools was randomly assigned as the external control group and all 18 students from these classes ($\underline{N}$ = 144) were allocated to the external control group. The intervention was not administered in these classes and no training or materials were provided to teachers of these classes. Because of the emphasis of this study on the comparison of internal control groups (where there is the possibility of diffusion of the experimental intervention) and the external control groups, analyses and discussion focus on these control groups and not the substantive interpretations of the intervention effects that are described in greater detail elsewhere (Craven, 1996).

Instrumentation

The SDQ-I (Marsh, 1990) was selected as the self-concept measure because it is widely regarded as the strongest multidimensional self-concept instrument for school-aged students (Byrne, 1996; Hattie, 1992; Wylie, 1989). The SDQ-I assesses three areas of academic self-concept (reading, mathematics and general school self-concept), four areas of nonacademic self-concept (physical ability, physical appearance, peer and parent relations) and includes a general self-scale. Three total scores consist of: academic self-concept (the average of reading, mathematics, and general school self-concepts), nonacademic self-concept (the average of physical, appearance, peer, and parent relations self-concept scales) and total self (the average of academic and nonacademic scales). Preadolescent children are asked to respond to 76 simple declarative sentences (e.g., “I'm good at mathematics”) with one of five responses: false; mostly false; sometimes true/sometimes false; mostly true; true. Because the diffusion effect is posited to generalize to different components of academic self-concept and because the initial selection of students and their assignment to groups were based on the total academic self-concept score, we based analyses on this score.

Intervention

Pretests were administered at the start of the academic year and the intervention was implemented during the next 14 weeks. The initial sample of 1557 pupils completed the SDQI, standardized achievement tests, and two other measures not associated with this aspect of the study (see Craven, 1996). The measures were administered by research assistants under the supervision of the first author according to testing procedures in the respective testing manuals. To examine the intervention effects, time 2 tests were administered 1- 3 weeks after the intervention.

Prior to administering the intervention, teachers of experimental classes attended one, 90-minute intervention training session. Teachers were instructed to praise 4 target children daily in specific subject areas (1 in mathematics, 1 in reading and 2 in both reading and mathematics) once each day. Teachers were instructed to deliver feedback daily during normal reading and mathematics lessons. Teachers were explicitly instructed to maintain their normal feedback for all students such that feedback associated with the intervention was in addition to normal feedback. The teacher-mediated intervention employed a combination of internally focused feedback and attributional feedback (see Craven, 1989; 1996; Craven, Marsh and

Debus, 1991). All activities were extensions of the procedures previously tested in the Craven et al. (1991) study.

During week 14 of the study, two external observers (senior research associates) who had observed the implementation of the intervention on 6 one hour occasions and teachers of experimental groups were asked to complete parallel questionnaires, commenting on and rating on a scale of 1 to 9 (1 - poor, 3 - below average, 5 - average, 7 - good, and 9 - excellent) the fidelity of the intervention implementation. Specifically teachers were asked to rate and comment on the item "My performance in ensuring only target pupils receive the teacher-administered intervention" and external observers rated and commented on the item "The teacher's performance in ensuring only target pupils receive the teacher-administered intervention". A total of 38 of 42 teachers completed the teacher self-rating form and 31 of 42 teachers were rated by observers who felt that they were able to make accurate ratings of the extent to which teachers had been able to focus the intervention on experimental students. These teacher self-ratings and ratings by external observers provided an indicator of the fidelity of implementation, a measure of the success of the teacher in focusing the intervention on target participants. These were collected as measures of fidelity of the implementation, but are also directly relevant to the evaluation of diffusion effects.

Statistical Analyses

In preliminary analyses, there were no significant pretest differences between internal within-class and external control groups on time 1 measures of academic self-concept or academic achievement. Diffusion effects in the teacher-mediated intervention were tested by contrasting the academic self-concept scores of the internal control group with the scores of the external diffusion control group at time 2. A multiple regression analysis was conducted with time 2 academic self-concept as the dependent variable. Covariates included: time 1 (pretest) scores for academic self-concept and academic achievement. Aptitude-treatment interactions -- the extent to which diffusion effects (operationalized as differences between the internal within-class controls and the external controls) varied as a function of initial academic self-concept -- were also evaluated in this multiple regression analysis (see Aiken and West, 1991). For purposes of these analyses, academic self-concept scores were standardized across the total group and achievement test scores were standardized across all students in the same year at school. Subsequent analyses were then used to determine the extent to which academic self-concepts of internal control students varied as a function of the teacher's success in focusing the intervention on target students.

Traditional multiple regression analyses such as those described above may be appropriate if there is no clustering effect (students within each class are no more similar to each other than to students from other classes), but this is unlikely in classroom research. When clustering effects do exist, tests of statistical significance are positively biased. Recent advances in multilevel modeling provide a means to evaluate whether there are such clustering effects and a more appropriate way to analyze the data whether or not there are clustering effects. A detailed presentation of the conduct of multilevel modeling (also referred to as hierarchical linear modeling) is available elsewhere (e.g., Bryk & Raudenbush, 1992; Goldstein, 1995). Particularly in social, organizational, and educational research, characteristics associated with individuals who are clustered within groups (e.g., students in classrooms, residents in neighborhoods, employees in companies) pose special problems related to appropriate levels of analysis, aggregation bias, heterogeneity of regression, and associated problems of model misspecification due to lack of independence between measurements at different levels. It is generally inappropriate to pool responses of individuals without regard to groups, and relations observed at one level may not bear any straightforward connection to relations observed at another.

Results

Diffusion Effects in Student Self-concepts

Comparison of time 2 self-concept scores for the internal and external control groups revealed main effects for group and an aptitude treatment interaction in which the size of this effect varied as a function of initial self-concept levels . Results based on the traditional (single level) multiple regression analyses and multilevel modeling analyses provide nearly identical estimates of the effects and their standard errors (see fixed effects in Table 1). Both analyses indicate significant main effects of pretest variables (prior academic self-concept and achievement) and significant main effects of (internal vs. external) groups in which scores are higher for the internal control group. This main effect of group, however, interacts with pretest academic self-concept. A preliminary multilevel model with no explanatory variables indicated that only a marginally significant portion of the variance could be explained by initial differences between classes (variance component = .092, SE = .047) and results in Table 1 indicate that the residual variance component after adding the explanatory variables is clearly nonsignificant (variance component = .010, SE = .022). This small clustering effect explains why results based on the two analyses are so similar.

Inspection of Figure 1 (considering only the solid lines representing the total internal and external control groups for now) demonstrates that these results support a diffusion effect; students in the internal control groups had higher academic self-concepts than did students in the external control groups. The size of this diffusion effect, however, varied with initial (pretest) levels of academic self-concept (T1 academic self-concept x internal interaction in Table 1). The diffusion effect was clearly evident for students with initially lower academic self-concepts but not for students with relatively higher academic self-concepts (i.e., high relative to this group of students with average and below average self-concepts). The nature of this interaction is consistent with the design of the intervention to enhance the self-concept of students with initially low self-concepts (see Craven, 1996, for a more detailed evaluation of this interaction effect).

Insert Table 1 and Figure 1 About Here

Focus: Measures of the Fidelity of Implementation

As with any experimental design resulting in significant differences between experimental and control groups – even those based on random assignment to groups, it is incumbent upon the researchers to support the construct validity of their interpretation of the cause of the group difference. To evaluate whether diffusion of the teacher-mediated intervention was the source of the group differences, data from teachers and external observers are examined. Initially we consider external observer and teacher self-ratings of focus (the extent to which teachers were able to focus the intervention exclusively on target students). Further data from comments by external observers and teachers on the implementation of this aspect of the intervention illuminate possible sources of the diffusion effect. Finally, these focus ratings are used to determine how the pattern of results varied as a function of the focus of the intervention.

Teacher responses. Teachers rated their ability to focus the intervention on experimental target students and not to other students in the class. Across self-ratings by all teachers, 13% were poor to below average (1-3 on a 9-point response scale), 47% were average (4-6 on the 9-point response scale), 39% were above average to excellent (7-9 on the 9-point response scale), and only 8% were excellent (9 on the 9-point response scale). The range of these self-assessment ratings suggests that teachers varied considerably in their ability or willingness to focus the intervention exclusively on target participants, supporting suggestions that the quality of implementation would vary from teacher-to-teacher.

Written comments by teachers also support the diffusion effect interpretation. One teacher noted that "I liked the intervention so much I used it with all my students". Other teachers expressed some difficulties isolating the intervention to nontarget students e.g.,: "I sometimes accidentally gave the intervention to someone not on the list", "I found it difficult to restrict who received the reinforcement". Others suggested that they thought it was beneficial to not solely isolate the intervention to target participants e.g.,: "I always naturally give other children positive reinforcement/feedback anyway so I didn't restrict it to just the target pupils". Some teachers suggested it was hard to focus on target children and that nontarget children might not receive praise when it was due e.g.,: "Hard to do as you felt guilty leaving others out", "I wanted to ensure the other children didn't feel any of the four children were getting special treatment", "It was difficult to cover each child and subject each day without leaving others out", "In a working classroom it is hard to not reinforce all workers if they are completing tasks credibly". Other comments suggested that teachers felt that nontarget class members were aware of the praise strategies being implemented e.g.,: "Awareness of class that some children were getting preferential treatment", "Some class members felt they also needed to be praised all the time".

External observer responses. External observers, based on classroom observations, also rated the ability of the teachers to focus the intervention on target experimental students. Observer ratings were systematically higher than teacher self-ratings, ranging from 4 to 9. Across all teachers they rated 29% as average (4-6 on a 9-point response scale), 71% as good to excellent (7-9 on a 9-point response scale), and only 13% of teachers as excellent in their ability to focus the intervention on target students. Ratings by external observers were more lenient than teacher self-assessments, but still indicate systematic variation in the ability of teachers to focus the intervention on target students. Written comments by observers also support this contention (e.g., "seemed to give the feedback to other students as well"; "was not prepared to praise only a small section of the class"). Hence, comments by the observers also indicated that some teachers did not focus the intervention on target students. Observers also reported that some teachers and even students felt that preferential treatment was being given to target subjects. For example, one observer noted that the teacher's performance in focusing the intervention "was to the point where the other students have felt left out even though the teacher has praised them in other ways". It was also noted that some teachers delivered the treatment in such a public manner that it was obvious to students e.g.,: "It is obvious which children are receiving the feedback", "It is quite obvious to me which are the participating students".

Statistical Analyses of Implementation Ratings

Teacher self-ratings and observer ratings of how well teachers were able to focus the intervention on experimental target students were positively correlated (r=.38, $p < .05$). In order to assess how student self-concept responses varied as a function of the focus of the implementation, a total focus score was obtained by averaging the nonmissing teacher self-ratings and observer ratings. Multiple regression analyses and multilevel analysis models were then used to determine the extent to which academic self-concepts of internal control students varied as a function of the teacher's success in focusing the intervention on target students. (These analyses did not include external control students because focus ratings were only relevant for internal control students and were not collected for the external control groups).

Results based on the traditional (single level) multiple regression approach and the multilevel modeling approach (Table 1) provide nearly identical results (see fixed effects Table 1, analysis of internal control group as a function of focus). Results for both analyses indicate a significant interaction effect (focus x T1 academic self-concept). Inspection of Figure 1 (considering only the broken lines for now), demonstrates that the difference between high and low focus groups are evident for students with relatively higher pretest academic self-concepts. These students with relatively higher self-concepts were more advantaged when the focus of the intervention was low (i.e., teachers were less able or willing to focus the intervention solely on experimental on target students; see Craven, 1996, for further discussion of this interaction effect).

It is also important to juxtapose the results based on the total internal and external groups (the solid lines in Figure 1) with the results for internal groups with a high and low focus (the broken lines in Figure 1). The function for the total internal control group, of course, falls midway between those based on groups with a high and low focus of implementation. The function for the internal control students with a low focus is systematically higher and roughly parallel with that of the external group. Consistent with our interpretation of the diffusion effect, this suggests that when teachers are not able to focus the intervention on target students, then students at all levels of prior academic self-concept benefit by the diffusion of the intervention. This finding, perhaps, constitutes the strongest support for a diffusion effect and demonstrates why it is important to simultaneously consider results from external control groups, internal control groups, and measures of the fidelity of the implementation. Interestingly, even when the teachers are able to focus the intervention on target students, the only nontarget students to be disadvantaged (relative to students in the low focus group) are those with relatively higher levels of initial self-concepts. When teachers are successful in focusing the intervention on experimental target students, nontarget students within the same class who might normally receive more praise and feedback (i.e., those with initially relatively high academic self-concepts) tend to have lower academic self-concepts (Figure 1).

Discussion

For purposes of this study, diffusion effects were operationally defined as students in the internal (within-class) control group having higher academic self-concepts as a result of a teacher-mediated intervention than students in the external (between-class) control groups. Even when researchers provide careful training to teachers and monitor the implementation, teachers are likely to differ in the fidelity of the implementation. The presence of diffusion effects in this study demonstrates that within-class control groups can be contaminated such that they also receive benefits of teacher-mediated interventions that are intended to be given only to experimental target students. Hence, in such studies the within-class control group provides a questionable basis of comparison for evaluating the effectiveness of the intervention. The external control group is, perhaps, a more effective control group in that students in this group are unlikely to be contaminated by the intervention. However, particularly for most classroom research based on modest sample sizes and a limited number of classrooms, the use of external control groups may not be a viable option.

Out results clearly supported a diffusion effect in that academic self-concepts were higher in the internal control group than in the external control group. Furthermore, consistent with the design of the intervention and its effects, the diffusion effects were limited primarily to students with initially lower levels of academic self-concept (Figure 1). The construct validity of the interpretation of this difference as a diffusion effect was supported by comments by both the teachers themselves and external observers. Furthermore, self-ratings by teachers of their success in focusing the intervention on experimental students varied widely and agreed reasonably well with parallel ratings based on responses by external observers. Particularly when teachers were unsuccessful in focusing the intervention on target children, academic self-concepts of nontarget students in the corresponding (low-focus) internal control groups were systematically higher for all levels of pretest academic self-concept than those in the external control group. Hence, the interpretation of a diffusion effect is supported by differences in student academic self-concepts, teacher and external observer comments and focus ratings, and differences in self-concepts as a function of teachers' success in focusing the intervention on experimental target students.

As emphasized by Plewis and Hurry (1998), much classroom intervention research is methodologically flawed in that the statistical analysis is not consistent with the focus of the intervention and the nature of educational data. Because educational data are hierarchically ordered (e.g., students are nested within classes), it is almost always appropriate to conduct multilevel analyses that that take into account this hierarchical data structure. Whereas it may be defensible to do analyses on class means, such analyses typically have insufficient power to identify potentially important intervention effects unless the number of classrooms is extremely large. Analyses at the individual student level are rarely justified in educational research in that tests of statistical significance are likely to be positively biased due to violations of the assumption of independence – that students are no more similar to students within the same class than to students in different classes. The major exception to this generalization is when such clustering effects are negligible. Interestingly, because the clustering effects were very small in the present study, our results based on the traditional multiple regressions (single level) and multilevel analyses were nearly identical. Even here, however, we needed to conduct the multilevel modeling in order to demonstrate that the clustering effects were small.

The focus of our research has been on diffusion effects on internal (within-class) control groups and associated biases in evaluating the effectiveness of interventions. There was clear evidence for a diffusion effect for the internal control groups with low focus. Thus, when teachers do not focus the intervention specifically on target students, students in the internal control group are likely to be benefited by the intervention. However, even when the focus of the intervention is high (i.e., fidelity of implementation is good), the results for the internal control groups were more complicated than anticipated. Differences between the high and low focus control groups varied with characteristics of the students as did differences between the high focus and external control groups (Figure 1). For example, differences between the high and low focus groups were larger for students with initially higher levels of self-concept. This suggests that diffusion (positive and constructive feedback from teachers) may be greater for these students who initially had relatively higher self-concepts. Furthermore, these students from high focus groups with initially higher self-concepts also appeared to be disadvantaged in comparison to even the external control group (Figure 1).

Although not anticipated, we offer several post hoc suggestions for why self-concepts of students with relatively high self-concepts might be lower in the high focus group than in the external control group. The results may be consistent with resentful demoralization hypothesized by Cook and Campbell (1979) in that students might feel resentful when similar students (target students with similar levels of academic self-concept) receive positive feedback and they do not. Similarly, this could represent a negative implicit reward (Sechrest, 1963; Sharpley, 1985) in which observing target students receiving praise results is a negative effect for students who do not receive praise even though their performance may be the same as target students. Furthermore, Sharpley (1985) emphasized that this effect is likely to be negative only when students previously have been praised for this behavior. From this perspective, it may be reasonable that this effect is negative in our study only for those students with relatively higher levels of academic self-concept. Alternatively, even though teachers were instructed to maintain their normal levels of praise for nontarget students, some may have been overzealous in not praising internal control students who might otherwise expect to be praised. Indeed, it may be realistic that teachers who see themselves as being highly focused in the administration of the intervention not only increase appropriate praise and effective feedback to intervention students but might also reduce normal levels of praise and feedback to other students – particularly those who might otherwise be most likely to receive it. This suggestion is also consistent with comments by teachers who felt guilty about withholding praise from internal control students, particularly those who most deserved it (also see Cook and Campbell, 1979, for related discussion of compensatory equalization). Furthermore, even if teachers maintained the same level of praise for nontarget students, these levels may seem to students to be less in comparison to the higher levels of praise received by targeted students. These alternative explanations are not mutually exclusive and may represent different perspectives of the same underlying phenomena. Indeed, Sharpley (1985) emphasized that the effects of implicit rewards in classroom settings are likely to be negative when teachers reduce the previous levels of reward experienced by nontarget students and that the effects are dependent upon worth of rewards as viewed by students.

It is important to emphasize that the results of the present investigation are idiosyncratic to particular characteristics of our study. It might be argued, for example, that diffusion effects are particularly likely in studies where the intervention is based on the administration of public praise and effective feedback and where the outcome variable is academic self-concept. The extent to which these results would generalize to different interventions and to different outcome variables is clearly beyond the scope of the present investigation. Also, because we did not have baseline patterns of reinforcement for our teachers prior to the introduction of the intervention, we can only infer how the introduction of the intervention changed these patterns of interaction. Instead, the focus of our research is to provide strong evidence that diffusion effects are possible when researchers rely on within-class internal control groups and to explore some of the likely implications of this effect as a bias to the valid interpretation of intervention effects. Although there has been considerable anecdotal reporting of diffusion-like effects for classroom intervention studies, the most appropriate evaluation of such effects requires an effective intervention, an internal within-class control group based on random assignment, an external control group based on random assignment, measures of the fidelity of implementation in the experimental classrooms, and appropriate statistical analyses to compare results for the internal and external control groups. Thus, it is not surprising that there has been little nonanecdotal support for diffusion effects based on true experimental designs. Hence, an important contribution of the present investigation is to demonstrate that under appropriate circumstances, the use of internal within-class control groups can result in diffusion effects that will bias interpretations of intervention effects.

More generally, the results have implications for the experimental design of classroom intervention studies. Particularly when there is a reasonable likelihood that the effects of an intervention may diffuse to other students within the same setting, sole reliance on an internal within-class control group is problematic. When teachers are unable or unwilling to focus the intervention on the target students (and instead, direct components of the intervention to control students), then there are likely to be substantial diffusion effects that bias the intervention effects. Also, because teachers are justifiably uncomfortable with the logistic, equity and ethical implications of introducing differential treatments that may deny their students access to an intervention that is seen to be beneficial, the fidelity of implementation of designs with internal control groups is likely to be variable. Furthermore, there may be many competing threats to the internal validity (e.g., diffusion effects, compensatory equalization, compensatory rivalry, resentful demoralization, implicit rewards and punishments) of interpretations of internal control groups comparisons that interact with characteristics to the study and the participants. Thus, it might be difficult to predict the size or even the direction of the cumulative effects of such biases. Hence, our over-riding message to researchers is to be wary of completely within-class experimental designs. Although we are not arguing that internal within-class designs are always biased or that there are not potential problems associated with external control groups, our results provide one clear example and rigorous tests of diffusion effects.

References

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis methods. Newbury Park, CA: Sage.

Byrne, B. M. (1996). Measuring self-concept across the life span: Issues and instrumentation. Washington, DC: American Psychological Association.

Cook, T. D., and Campbell, D. T. (1979). Quasi-experimentation. Design and analysis issues for field settings. Chicago: Rand McNally Publishing Company.

Clarke, R. V. G., & Cornish, D. B. (1972). The controlled trial in institutional research -- Paradigm or pitfall for penal evaluators? Home Office Research Studies, 15. London: HMSO.

Cooper, H. M. (1977). Controlling personal rewards: Professional teachers' differential use of feedback and the effects of feedback on the student's motivation to perform. Journal of Educational Psychology, 69, 419-427.

Craven, R. G. (1989). An examination of self-concept: The interrelationship of teachers', parents and children's perceptions of self-concept, and their influence in enhancing self-concept. Unpublished B.A. Honours Thesis, University of Sydney, Australia.

Craven, R. G. (1996). Enhancing Academic Self-Concept: A Large-Scale Longitudinal Study in an Educational Setting. Doctoral thesis submitted to the University of Sydney. USA: UMI.

Craven, R. G., Marsh, H. W., and Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on the enhancement of academic self-concept. Journal of Educational Psychology, 83, 17-27.

Goldstein, H. (1995). Multilevel statistical models (2nd ed.). London: Arnold.

Good, T. L., and Brophy, J. E. (1974). Changing teacher and student behavior: An empirical investigation. Journal of Educational Psychology, 66, 390-405.

Hattie, J. A. (1992). Self-Concept . Hillsdale NJ.: Lawrence Erlbaum Associates.

Marsh, H. W. (1990b). The Self Description Questionnaire-I: SDQ-I manual. Sydney: University of Western Sydney.

Plewis, I., & Hurry, J. (1998). A multilevel perspective on the design and analysis of intervention studies. Educational Research and Evaluation, 4, 13-26.

Sechrest, L. (1963). Implicit reinforcement of responses. Journal of Educational Psychology, 54, 197-201.

Sharpley, C. F. (1985). Implicit rewards in the classroom. Contemporary Educational Psychology, 10, 349-368.

Withall, J. (1956). An objective measurement of a teacher's classroom interactions. Journal of Educational Psychology, 47, 203-212.

Wylie, R. C. (1989). Measures of self-concept. Lincoln: University of Nebraska Press.

Table 1

Time 2 Academic Self-concept for Internal and External Comparison Groups and as a Function of Focus in the Internal Comparison Group

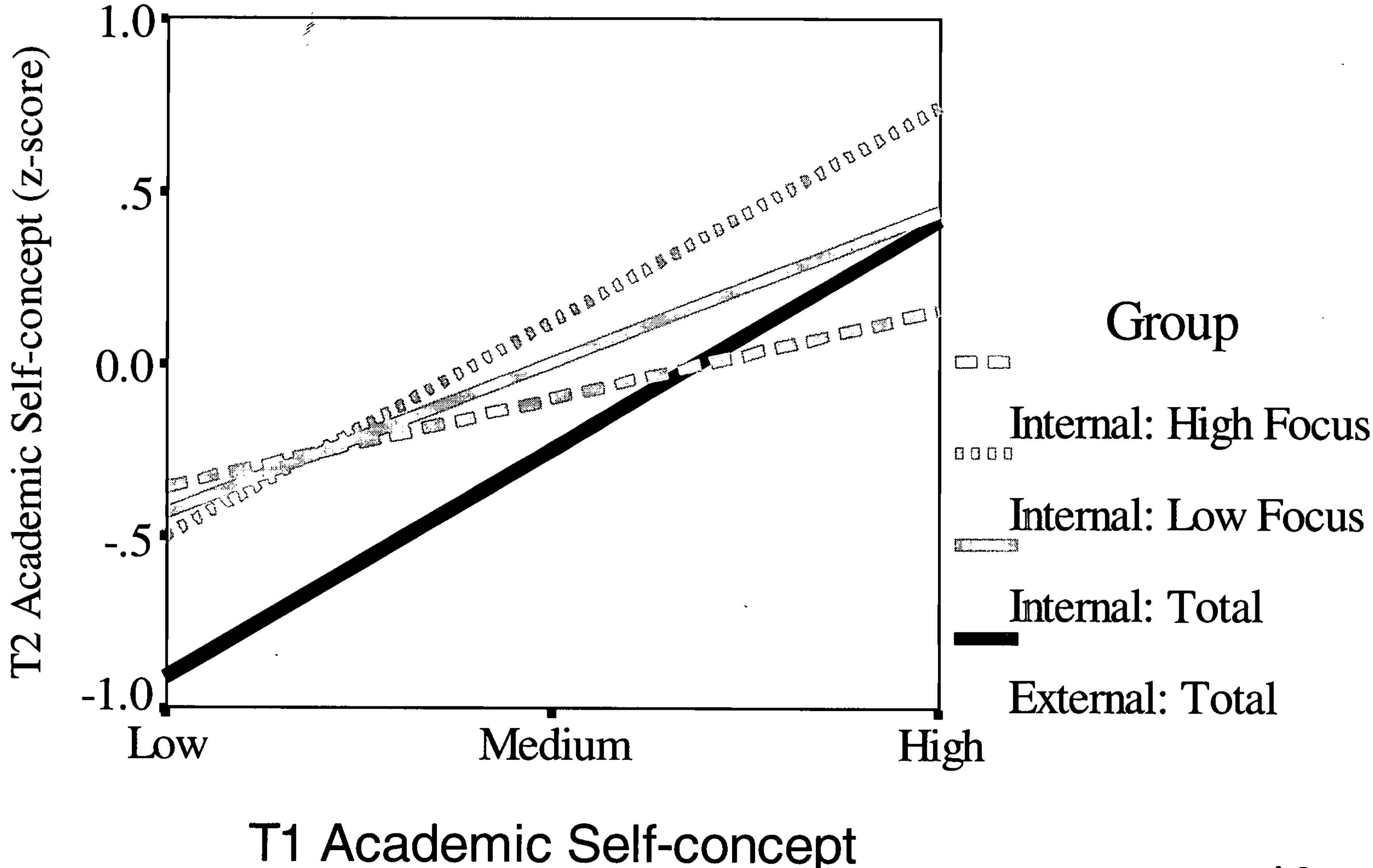
	Multiple Regression				Multilevel Modeling	
	B	SE B	Beta	Partial	Parm	SE
Internal vs. External Control Groups						
Fixed Effects						
T1 ASC	.51*	.05	.48	.48	.51*	.05
Internal/External	.08*	.03	.12	.14	.08*	.03
T1 ASC x Internal	-.07*	.04	-.09	-.11	-.07*	.04
T1 Achievement	.13*	.05	.12	.13	.13*	.05
(CONSTANT)	-.08	.05			-.08	.05
Random Effects						
Student (level 1)					.72**	.06
Class (level 2)					.01	.02
Effect of Focus (Internal Group only)						
T1 ASC	.44*	.05	.48	.46	.43*	.06
Focus	-.08	.06	-.08	.07	-.09	.07
T1 ASC x Focus	-.12*	.06	-.12	-.09	-.15**	.06
T1 Achievement	.07	.07	.06	-.14	.11	.07
(CONSTANT)	.01	.05			.00	.07
Random Effects						
Class (level 2)					.07	.04
Student (level 1)					.61**	.06

Note. T1 ASC = Time 1 academic self-concept. Internal = Internal vs. External contrast (positive coefficients indicate higher scores for the internal control group).B = unstandardized beta weight. SE B = standard error of the unstandardized beta weight. Beta = standardized beta weight. Partial = partial correlation. Parm = Parameter estimates from multilevel modeling. SE = Standard errors from parameter estimates from multilevel modeling. Also, see Figure 1 for a graph of the effects.

* $p < .05$; ** $p < .01$.

BEST COPY AVAILABLE

Figure 1.Academic self-concept at T2 as a function of Group and T1 Academic Self-concept (Low = +1 SD, Medium = mean, High = +1 SD). Two groups consist of the total internal and external control groups. The Internal Control Group is also evaluated for cases where the focus of the intervention was either high (+1.5 SD) or low (-1.5 SD)



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Diffusion Effect 1 Teacher-Administered Self-Concept Interventions: Do Diffusion Effects Exist.?

Author(s): Rhonda G. Craven

Corporate Source: Self-concept Enhancement and Learning Facilitation Research Centre

Publication Date: 24 April 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: R. Craven

Printed Name/Position/Title: Dr. Rhonda Craven

Organization/Address: SELF Research Centre University of Western Sydney PO Box 555 Campbelltown Australia 2560

Telephone: +61 2 9772 6557

FAX: +61 2 9772 6432

E-Mail Address: r.craven@uws.edu.au

Date: 10/4/2000

ERIC